DOCUMENT RESUME

ED 274 690 TM 860 558

AUTHOR Baldwin, Beatrice
TITLE Some Cautionary Notes on the Specification and Interpretation of LISREL-type Structural Equation Models.
PUB DATE Apr 86
NOTE 39p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Analysis of Covariance; *Computer Software; *Equations (Mathematics); Error of Measurement; *Mathematical Models; *Measurement Techniques; *Path Analysis; Research Methodology; Social Science Research; Statistical Studies
IDENTIFIERS *LISREL Computer Program

ABSTRACT

LISREL-type structural equation modeling is a powerful statistical technique that seems appropriate for social science variables which are complex and difficult to measure. The literature on the specification, estimation, and testing of such models is voluminous. The greatest proportion of this literature, however, focuses on the technical aspects of LISREL, and many are far advanced in statistical sophistication so as to be of little or no help to the novice modeler. This paper presents a "philosophical" approach to structural equation modeling and examines some of the basic concepts and assumptions underlying the formulation of these models. The rationales for the structural equation model and the measurement model are outlined, and the foundations of specification and interpretation clarified. As part of the SPSS-X package, LISREL will be readily available to many applied researchers. With this heightened availability of a complicated statistical procedure will undoubtedly come many misuses and abuses. Social scientists should become familiar with the theoretical premises of the method. (Author/JAZ)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED274690

Some Cautionary Notes on the Specification and Interpretation of LISREL-type Structural Equation Models

Beatrice Baldwin
Louisiana State University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

B. Baldwin

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, 1986.

BEST COPY AVAILABLE

TM 860 558

Abstract

Use of LISREL-type structural equation modeling has become more widespread in the social sciences, and the literature on the specification, estimation, and testing of such models is voluminous. The greatest proportion of this literature, however, focuses on the technical aspects of LISREL, and many are far advanced in statistical sophistication so as to be of little or no help to the novice modeler. This paper presents a "philosophical" approach to structural equation modeling and examines some of the basic concepts and assumptions underlying the formulation of these models. The rationales for the structural equation model and the measurement model are outlined, and the foundations of specification and interpretation clarified. As part of the SPSS-X package, LISREL will be readily available to many applied researchers. With this heightened availability of a complicated statistical procedure will undoubtedly come many misuses and abuses. Social scientists should become familiar with the theoretical premises of the method.

Some Cautionary Notes on the Specification and Interpretation of LISREL-type Structural Equation Modeling

Recent advances in the computerized application of sophisticated statistical methodologies have enabled social scientists in economics, psychology, education, sociology, and related disciplines to utilize multivariate analysis techniques previously beyond the scope of these fields. In particular, the linear structural relations model (LISREL) defined and developed by Joreskog and Sorbom (1984) has given rise to a new emphasis on covariance structure analysis. The field of covariance structure analysis actually includes several related topics: (1) path analysis (2) factor analysis and (3) structural equation modeling which combines the analytical benefits of the first two. Technical literature on the specification, identification, estimation, and testing of linear models abounds. An overview of the detailed methodology is beyond the scope of this paper, but introductions can be found in Bentler (1980), Joreskog & Sorbom (1984), and Lomax (1982). The social scientist who possesses a solid background in statistical and mathematical theory has benefitted most from the literature on the LISREL model, since the greatest proportion concentrates on technical aspects and state-of-the-art applications. The novice modeler, however, who

possesses limited knowledge of intricate statistical theory may find the bulk of the literature unsuitable as a general introduction to the use of the LISREL computer program. Furthermore, theoretical and philosophical considerations of LISREL methodology, its assumptions, and its interpretations are few. The purpose of the following paper is two-fold: (1) for the novice, an introduction to the concepts and assumptions basic to LISREL structural equation modeling, and (2) for the researcher already somewhat familiar with the technical aspects of covariance structure analysis, a clarification of those theoretical aspects of LISREL modeling which have been most misunderstood.

The LISREL Model- An Elementary Example

Simply put, covariance structure analysis involves the analytic mathematical breakdown of the covariances between measured variables into estimates of the strength of relationship between latent variables. Consider the simple model represented in Figure 1. The ellipses ξ_1, ξ_2, and η_1 represent hypothetical constructs or latent variables. Such latent variables cannot be measured directly, and so are sometimes referred to as unobserved variables. For instance, as a subset of a more complicated model, Baldwin (1984) considers the role-modeling aspects of mother-daughter relationships and hypothesizes that the extent of a mother's influence on her

daughter's academic behavior (ξ_1) and the nature of the daughter's sex role orientation (ξ_2) will influence the educational and occupational aspirations of the daughter (η_1) and consequently her achievement (η_2). These four variables are latent variables in that the natures of the constructs are concealed and cannot be observed or quantified directly.

Insert Figure 1 About Here

The boxes labeled X_1, X_2, X_3, X_4, Y_1, Y_2, Y_3, and Y_4, represent indicator or manifest variables. Such variables are actual measures and collectively serve to represent the unoberved latent variable. For instance, educational/occupational aspiration is a latent construct which cannot be observed and measured itself. However, by measuring the number of years of schooling a daughter expects to complete and by determining occupational ambition according to a socioeconomic index, one can obtain a rough "quantification" of a concept that could not be estimated directly.

The relationships of the latent variables (the ξ's and the η's) are represented by structural equations of the form

$$\eta = \beta\eta + \Gamma\xi + \zeta$$

and the paths between the variables (the γ's and the β's) are

called structure coefficients. That part of the model which defines the relationships between the latent exogenous (independent) and endogenous (dependent) variables is called the structural equation model. For this particular example, there are two structural equations

$$\eta_1 = \gamma_{11}\xi_1 + \gamma_{12}\xi_2 + \zeta_1$$
$$\eta_2 = \beta_{21}\eta_1 + \zeta_2$$

and three structure coefficients γ_{11}, γ_{12}, and β_{21}.

The relationships between the manifest variables and their corresponding latent variables are represented by the equations of the measurement model:

$$X_1 = \lambda_{11}\xi_1 + \delta_1$$
$$X_2 = \lambda_{21}\xi_1 + \delta_2$$
$$X_3 = \lambda_{32}\xi_2 + \delta_3$$
$$X_4 = \lambda_{42}\xi_2 + \delta_4$$
$$Y_1 = \lambda_{11}\eta_1 + \epsilon_1$$
$$Y_2 = \lambda_{21}\eta_1 + \epsilon_2$$
$$Y_3 = \lambda_{32}\eta_2 + \epsilon_3$$
$$Y_4 = \lambda_{42}\eta_2 + \epsilon_4$$

There are eight equations and eight factor loadings (the λ's). Each indicator variable only loads on one latent variable.

Together, the structural equation model and the measurement model make up the LISREL model which can be analyzed via the LISREL computer program. As can be seen by the example, the

LISREL model can be utilized in behavorial research sitauations characterized by the presence of latent variables. The power of the LISREL computer program lies not only in its ability to calculate parameter estimates, i. e., the structure coefficients, factor loadings, measurement and equation error terms, and latent variable correlations, but also in its capabilities to compute standard errors, goodness-of-fit indices, standardized solutions, and indices to suggest possible modifications to the model. All of these potential analytical tools make LISREL a formidable statistical technique, a technique which can be of tremendous exploratory or confirmatory value when used in accordance with the theoretical principles on which it is based.

The Rationale of Structural Equation Modeling

Structural equation modeling, linear structural relations, covariance structure analysis, and simultaneous equation modeling are synonymous terms for the same methodological approach. All of these terms can be classed under the more general rubric of causal modeling. In a very loose sense, causal modeling infers "causation." That is, a change in some variable X causes an effect-- a change in another variable Y. The cause and effect is direct and in many cases unidirectional or recursive. In the LISREL example, it was hypothesized that the nature of a daughter's sex role orientation would influence her educational and occupational aspirations; that is, a change in sex role

orientation would directly cause a change in educational/ occupational aspirations.

Sophisticated programs like LISREL IV have increased the rigor by which researchers can analyze data but at the same time present severe interpretational problems when making causal inferences about hypothetical constructs (Cliff, 1983). As the levels of complexity and sophistication associated with programs like LISREL rise, it becomes relatively easy for the researcher to lose sight of one of the most fundamental principles of research-- that inviolable caveat, "Correlation does not infer causality." Covariation among variables may or may not represent a cause; the covariation may be due to spurious correlation, or there may be significant intervening variables masking the directness of the effect. The most satisfactory method for inferring causality is still the true experiment, whereby the researcher maintains active control of the variables. However, in social science research the true experiment is a rarity. Rather the researcher usually relies on prior experience, personal observation, and the prevailing theories of the discipline. In many instances though, the experience may be scant, the observation tainted, and the theory contradictory. In many instances causal inference of a dubious nature may result in poor modeling practices and incorrect parameter estimates. In as much as the structural parameters reflect the strength of causal relationships, poor models reinforce poor theory and do little to

advance fundamental knowledge in the field.

James, Mulaik, & Brett (1982) make a distinction between a functional relation which is deterministic and a functional relation which is probabilistic. For example, the linear equation

$$Y = \beta_{yx_1} X_1 + \beta_{yx_2} X_2$$

is deterministic because X_1 and X_2 completely determine Y. However, in any one social science research question, the number of possible functional relationships between all potential variables is for all practical purposes infinite. The deterministic notion then of accurately specifying all components of a causal relationship is unreasonable. Instead one must take a probabilistic view of causality and specify a linear relationship by

$$Y = \beta_{yx_1} X_1 + \beta_{yx_2} X_2 + e$$

wher e is a random error or <u>disturbance term</u> which accounts for all of the unknown influences not specified by X_1 or X_2.

The problematic distinction then comes between what variables are explicitly included in the model and what variables are represented by the component <u>e</u>. James et al. refers to this distinction as the difference between a <u>measured relevant cause</u> and an <u>unmeasured relevant cause</u>. The failure to include all relevant causes in a structural equation model results in biased parameter estimates and erroneous causal conclusions. A relevant cause is defined as a variable that:

(a) has a non-minor, direct influence on an effect,

(b) is stable,

(c) is related to at least one other cause included explicitly in a functional equation, and

(d) makes a unique contribution to a functional equation... (James et al., p. 23).

The inclusion of all relevant causes in a functional equation guarantees that the disturbance term $\underline{e}$ will include only those causes of Y that are minor, indirect, unstable, independent, and random. Therefore, a researcher must think of a causal model as a subsystem of a larger theoretical system, but that subsystem must be "closed" or self-contained. In practicality, no model can ever be truly self-contained; behavorial theory is too complex for such gross over-simplification. Careful consideration and thoughtfulness in specifying the structural model, however, can result in better models and more meaningful results.

The existence of unmeasured relevant causes is a type of specification error which results in covariation between the explicit causes (X_1 and X_2 in the linear example) and the disturbance term $\underline{e}$. Such covariation is a violation of the assumptions of the general linear model, and the results may be biased parameter estimates and false conclusions based on improper modeling. Since all relevant variables cannot be practically entered into the model, the problem then reduces to

constructing the model such that all known relevant causes are included while recognizing that such a procedure is guaranteed to be theoretically incomplete (Blalock, 1982).

James (1980) outlines a series of decision steps which can be used to determine if a cause is relevant or not. These decision steps are somewhat subjective and untestable in the verification sense, yet may provide a basis for model-building. The first step consists of identifying all known major and moderate causes of the dependent variable. The researcher should decide whether any unspecified cause correlates with any cause already in the model. At this point, James encourages a decision based on the absolute value of the correlation; a low correlation (- .20 to .20 is suggested) would allow the researcher to disregard the unspecified cause as its inclusion in the model wpuld probably not affect parameter estimates. A moderate to high correlation however warrants a model alteration if the variable makes a unique contribution to the dependent variable. The unique contribution can be viewed from the question of redundancy, i.e., if there is a very high correlation between two variables, then a multicollinearity problem will result. The variables are nearly one and the same, and the variance contributed by the second is not unique. Correlations between unspecified and specified causes may be estimated by references to past research, by theory, or by exploratory study. Without the inclusion of a relevant cause, serious bias in parameter estimates may result,

and the model cannot be considered as representative of reality.

As we have seen, the judicious selection of variables based on the seemingly true nature of causality and the approximation of a self-contained system will result in the specification of improved models. Two other considerations must also enter into the model-building process: (1) the specification of causal order, and (2) the specification of causal direction. Causal order is indicative of the temporal sequence of cause and effect; that is, we assume that if X causes Y, X occurred first and then Y occurred as a result of X. It also then natural to think of a specific time lapse between the occurrences of X and Y, although in some cases the elapsed time may be so insignificant so as to appear that X and Y are simultaneous. In structural equation models, it is not necessary to specify the causal interval, but a causal order must be specified. The arrangement of the variables in the graphic representation of the model relates the causal order of the latent constructs. So in the mother-daughter example, a change in the mother's influence in academics precedes a change in educational/occupational aspirations.

As with the selection of variables, the specification of causal order can be quite difficult, especially in the absence of recognized guiding theory. Insignificant or unobservable causal intervals may lead the researcher to misspecify causal order by inferring simultaneity when it does not exist or by reversing the true causal order.

Structural equation models likewise require the specification of causal direction. So far, the mother-daughter example has only considered a unidirectional flow of causality. A model which is unidirectional is termed recursive while a model that considers reciprocal causation is called non-recursive. Suppose that the model in Figure 1 were modified to include a reciprocal relationship as in Figure 2 (Of course, there must be some theoretical justification for such a modification. In this case let us simply assume that theory suggests a possibility.) Although the graphic representation of the model does not appear to be significantly different, the structural equations by which the parameters are estimated are quite different. Now there is only one exogenous (independent) variable ξ_1 and three endogenous (dependent) variables η_1, η_2, and η_3. The structural equations are

$$\eta_1 = \beta_{12}\eta_2 + \zeta_1$$
$$\eta_2 = \beta_{21}\eta_1 + \gamma_{21}\xi_1 + \zeta_2$$
$$\eta_3 = \beta_{32}\eta_2 + \zeta_3$$

and four structure coefficients γ_{21}, β_{12}, β_{21}, and β_{32} will be estimated. There are also resulting changes in the equations for the measurement model.

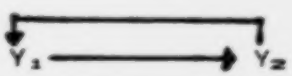
Insert Figure 2 About Here

Note that in this non-recursive example, causal order is

still maintained, i.e., a change in sex role orientation still precedes a change in educational/occupational aspirations. In other models, however, causal order may be irrelevant, so that reciprocal causation may be depicted as

in which case either X or Y may "initiate" the causal action. The effect of mutual causation is a dynamic system, and it is assumed that such reciprocity reaches a theoretical state of equilibrium (Namboordiri, Carter, & Blalock, 1975). The causal interval is insignificant and in most cases may be regaredd as approaching zero.

A differentiation must be made between reciprocal causation and cyclical causation which is an influential factor in some types of analysis. Cyclical causation may be depicted as



where Y_1 causes Y_2 which in turn influences Y_1 through a feedback loop, the difference being that a known time interval exists between the occurrences of Y_1 and Y_2. In actuality, the model would be specified as

$$Y_{1t_1} \longrightarrow Y_{2t_2} \longrightarrow Y_{1t_3} \longrightarrow Y_{2t_4} \longrightarrow Y_{1t_5}$$

where t specifies a time indicator. The number of occurrences of

Y_1 and Y_2 would be determined by the exact form of the design and the resultant data collection. This type of cyclical causation then would be treated as recursive since all causal effects are unidirectional (Strotz & Wold, 1971).

In the consideration of the rationale of the structural equation model, one must consider another important prerequisite for the specification of causal models-- the issue of moderating variables or "contextual boundaries" (James et al., 1982). Such variables are contingent upon the types of subjects and the environment of the analysis and severely limit the extent of generalization which is, of course, one of the primary aims of social science research. For example, the modeling of mother-daughter relationships as depicted in Figure 1 is limited to such relationships and could not be generalized to father-daughter, mother-son, or father-son relationships unless prevalent theory guided one toward that conclusion. The uniqueness of sex role related experiences would make such a conclusion improbable however, and alternative models might be specified and estimated for the alternative relationships. Even in the event that such a model might be conjectured for each of the parent-child subgroups, it would be beneficial to conduct each analysis separately and test the derived parameters for an interaction effect. Such interaction effects are common in studies involving variables of classification, i.e., sex, SES, race, age, or any other variable which could theoretically divide

a sample into differing subgroups. Failure to specify contextual boundaries and account for varying subgroups via differential model specification may make conclusions less meaningful and generalization impossible.

In conclusion, when structural equation models are used in theory-building, the emphasis is on defining the model so as to correspond to the perceived true nature of causality. In other words, one would wish to specify as exactly as possible how a change in one variable affects another variable. In any examination of causal relationships, one would want to know (1) if a relationship actually exists (that is, there exists a non-zero correlation); (2) if a supposed causal relationship is spurious; (3) if the order and directionality of causality can be specified; (4) if the causation is direct or indirect; (5) if reciprocal causation exists; and (6) if some moderating variable may be producing an interaction effect.

The reasons for using latent variables in a structural equation model seem self-evident: latent variables are abstractions-- hypothetical constructs which cannot be directly observed or measured. Since direct measurement is impossible, obervable indicator variables must be used as representations of the constructs. However, the researcher must guard against oversimplification and thoughtless modeling which has little regard for substantive theory. Both methodological and theoretical assumptions must be met before the model can

reproduce the strength and nature of the effects. Structural equation modeling is explicit and quantitative; it allows us to "explain" how variables relate to each other. This type of modeling also has simulation power in that it permits the researcher to consider an entire working subsystem and its character (Heise, 1969).

The Rationale of Measurement Modeling

There is a functional distinction between the philosophical bases of the measurement model and the structural model. The measurement model is not concerned with the notion of causality; the fundamental significance of the model stems from the correlations that indicate the relative ability of the known estimators to predict a value for a theoretical construct. The prime criteria for the evaluation of measurement models are efficiency and accuracy. Structural equation models are usually more hypothetical and therefore more complex in one sense. The specification must be theoretically correct, and it must be quantitaively valid so as to mathematically reproduce causation. Measurement models, however, are more pragmatic and may be less nebulous. The primary concerns are the reliability and validity of measurements.

Measurement models represent attempts to define latent constructs through manifest and observable variables. When only one indicator variable represents a latent variable, the

indicator variable serves as a surrogate for that which cannot be directly estimated, and perfect reliability (i.e., no measurement error) is assumed. If all latent variables in a model are represented by only one corresponding indicator variable, then no measurement model is necessary. The relationships between the manifest variables are represented by the structural equations, and the two-part LISREL model is reduced to a path analysis model.

In this case, the reliability coefficient for each measure is assumed to be 1.00, although in actuality a "high" reliability satisfies the operationalization of latent variables without much distortion of parameter estimates (Joreskog, 1979). Yet practicality in the social sciences dictates that perfect measurement is an impossibility, and consistent high reliability of measurement is an improbability. Tests and measures designed by human beings to measure latent characteristics of other human beings will be fallible by their very nature and subject to varying proportions of error.

While a degree of error is always expected, the magnitude of that error is of chief concern. As stated previously, one manifest variable vcould be selected to represent a latent construct, but perfect reliability must be assumed and a corresponding loss of accuracy is inevitable. A second sclution is to combine a number of measures into one variable which serves as an index. This method may preserve some information and

improve accuracy, but will not be as helpful as using several independent indicators. Use of multiple indicators increases the number of testable predictions, increases validity, and guides respecification of the model if an initial analysis proves unacceptable (Sullivan, 1971).

The use of separate indicators however may pose problems in complex models. Large numbers of multiple indicators become unmanageable in practical applications of the LISREL computer program. Thus by restricting the number of indicators or by using composites, one tends to lose the advantage of multiple indicators when, in the analysis of complex models, these advantages are needed most. The consideration of reliability becomes essential when selecting measurements in this case.

The problems of indicator reliability surface frequently in LISREL model specification. If necessary, one could use the available measure as is and simply hope that the resulting parameter estimates are not overly biased. Use of multiple measures, however, increases overall reliabilty even if the individual reliabilities are considerably less than perfect. A third possibility is to correct the correlation coefficients for unreliability and then to use the disattenuated coefficients as input to the LISREL program. One problem with this approach is that the reliability of the instrument must be known which in some instances is not the case.

Lomax (1986) demonstrates that the notion of measurement

error for LISREL models is somewhat different from the common notion of measurement error derived from classical test theory. In classical test theory

Total observed variation = True variation + Error variation
(common + specific)

where the true variation consists of common variation and specific variation which is non-random, systematic, and due to the nature of the particular variable. In contrast factor analysis (the LISREL measurement model) assumes that

Total observed variation = Common variation + Unique variation.
(specific + error)

Unique variation in this model consists of specific variation plus error which is random, unsystematic and due to unreliability. Thus in order to examine the effects of unreliability in LISREL models one must decompose the unique variation into its two components. Lomax's study of the effects of unreliability presents several important points which must be considered in the formulation of measurement models: (1) LISREL "measurement error" estimates may be deceptive as these estimates actually consist of specific variation and error due to unreliability; (2) if reliability for the indicators of one

latent variable are reduced relative to other indicator reliabilities, the resultant estimates for the structure coefficients will be decreased and therefore biased; (3)standardized parameter estimates do not seem to be as affected by unreliability as unstandardized maximum likelihood estimates; and (4) if indicator reliabilities are known and are less than .80, the correlations should be corrected before being used as input. Thus the situation of using a singular indicator variableis far from ideal. Using multiple indicators is always preferable and using disattenuated correlations in situations of low reliability may also be desireable.

Substantive knowledge of possible indicator flaws may aid the researcher in determining approximate levels of reliability if the exact reliability coefficient is unknown. Knowledge of ceiling or floor effects, test administration problems, systematic measurement error, and other such sources of possible bias may provide clues to possible measurement model alterations if an initial analysis shows unreliability of some indicators. Most commonly, an inspection of the residual matrices may reveal some unually high residual values relative to the other residual estimates. Such high estimates may be indicative of measurement unreliability. Costner & Schoenberg (1973) suggest an examination of residuals followed by the fitting of alternative models. Progressive deletions of suspected unreliable indicators can permit a diagnostic analysis of the effects of

unreliability on parameter estimates. Such a procedure could be quite tedious and time-consuming but may prove indispensible if unreliability problems are suspected and other knowledge of indicator accuracy is scant.

Cliff (1983) draws attention to the problems of reliability and validity in discussion of the "noministic fallacy." That is, naming a property or trait does not imply understanding of that property or even naming it correctly. There is a basic contextual gap between the hypothetical composition of the latent variable and the realistic composition of its indicators. The size of this contextual difference depends not only on the unreliability problem, the problem of consistency, but also on the invalidity problem, the problem of accurate representation. Special attention must be given to the appropriate selection of variables so as to maximize validity in an efficient manner. The addition of multiple indicators may increase reliability but will not increase validity if the validity of each individual indicator is doubtful. Validity will however be strengthened if the nature of each measure does accurately correspond to the nature of the construct. Usually correlational data is only suggestive of the underlying construct. In other words, high correlations between associated measures may aid in selecting which indicators should be utilized, but cannot guarantee the construct validity of the latent variable. The primary question of validity must be addressed by substantive theory, critical

knowledge of the indicator, and precision in defining the latent variable. The definition and interpretation of latent variables only becomes less uncertain when theory guides selection and when the individual indicators are valid and reliable.

Most variables of interest to social scientists are latent variables, conceptualized to be continuous at the latent variable level. At the indicator level, however, a variable may be ordinal and discrete. Yet uses of multivariate parametric statistics, including LISREL, usually require assumptions of multivariate normality. LISREL models then are formulated in mathematical terms which take interval level measurement for granted. The propositions of these models express values of variables as functions of other variables. Wilson (1971) states that when measurements are ordinal, meaningful estimates are usually impossible to obtain. Ordinal variables can be viewed as sets of mutually exclusive categories such that the categories can be put on a continuum. For instance, an attitude scale may have categories ranging from "strongly agree" to "strongly disagree." Any sequence of values can be assigned to the corresponding categories. The values are essentially arbitrary. Only the order of the values as reflects the order of the categories is significant. According to Wilson, in certain circumstances some models using ordinal data may be used to derive causal inferences but inferences are weak, and the model itself is usually faulty.

In contrast, Borgatta & Bohrnstedt (1981) argue that it intuitively makes sense that the bulk of most observations should lie close to the mean with relatively fewer cases at the extremes. Therefore it should be possible in most cases to assume that the distribution of most indicator variables should approximate normality.

Social measurement is usually crude in comparison to the sophisticated measurement possible in the physical and natural sciences. Measures are likely to be imperfectly continuous and non-interval in the sense that derived categorical intervals are not truly equal. However, if a measure is developed carefully, one should be able to assume a monotonic relationship between the imperfect indicator variable and the continuous latent variable. That is, a difference between two points on the manifest variable scale should reflect a coarse approximation to the corresponding difference on the latent variable scale. Measurement error therefore could be defined as the difference between the observed reality of the indicator variable scale and the unmeasureable reality of the latent variable scale. This measurement error will have an effect on the estimation of the model.

Furthermore, Borgatta & Bohrnstedt point out that the concept of ordinal data in most social science research does not reflect the true concept of ordinal position. A member of a population is not truly numerically "ranked" relative to every other member of that population. Ordinal scaling usually results from a

transformation in which interval information is lost. Some researchers contend that the loss of information is so significant as to render the use of ordinal data useless--insensitive to the basic laws of mathematics and algebra and consequently to the axioms of parametric statistics. If this is true then computing a tau or Wilcoxon test should make no sense as the computations require the mathematical addition of ranks. In reality, these tests transform ordinal ranks into a form of the interval scale such that a distance of one is between each value. That this unit distance does not perfectly correspond to a unit of distance on the latent interval scale is of less consequence than might be thought. One must reflect on the totality of the manifest variable distribution and its approximation of the latent variable distribution.

Lomax (1983) offers a guideline which may be useful if the researcher feels that the ordinal variables being used are approximating a normal distribution; if the number of categories for the variable is greater than or equal to 4 and if skewness does not exceed 2.0, then the underlying distribution is approximately normal and resultant parameter estimates will not be significantly affected. Data-gathering on any level should maximize the information being collected given the limiting circumstances under which measurements will be made. Thus decisions concerning ordinal vs. interval scaling must be made on a pragmatic basis, weighing the cost of presumed accuracy in

balance with the advantages of efficiency.

The Interpretation of LISREL Models

If the specification of the structural and measurement models can be considered correct-- that is, that the equations approximate reality given the limitations of social science-- the problems of interpretation may still affect results. Technical aspects of interpretation can be found in numerous papers (for example, see Bentler & Bonetett (1980) for guidelines on analysis of goodness-of-fit, and the LISREL VI User's Guide (Joreskog & Sorbom, 1984) for information on modification indices, goodness-of-fit indices, and related technical analysis). Let us consider instead some philosophical viewpoints in regard to interpretation.

One principle of scientific inference which may be violated is the notion that the data do not confirm the model even if the model does have a "good fit" as indicated by the chi-square test statistic. The researcher can only fail to disconfirm the model given the unique theoretical rationale which guided the formulation of the model (alternate theoretical explanations may be just as feasible) and given the specific data set used for analysis (conclusions may be sample-specific). In actuality, an infinite number of different models may be equally plausible in the theoretical and mathematical sense. Goodness-of-fit only implies that the particular model under consideration has the mathematical capability of reproducing the sample covariance

matrix.

"Confirmation" in the LISREL sense means that the predictions regarding the covariances of the indicator variables are consistent with the empirically derived covariances. If the model is disconfirmed or rejected, then the predictions regarding variable relationships are inconsistent with the sample data (James et al., 1982). The confirmation of predictions seems to imply corroboration, but this corroborations is meaningful only in a very limited sense.

In the development of a theoretical model of structural equations, the inclusion of a causal path between variables indicates that the expected value of that path should be significantly different from zero. Failure to include such a path in the model then indicates a prediction of zero for the corresponding structural parameter. By conducting tests of significance on the estimates, one can ascertain whether a prediction is consistent with the data. It may be possible that some path coefficients are non-significant even with careful attention to theory and the statistical meaningfulness of a large sample. One might ask how serious such misspecifications may be. The answer is conditional and depends on the model, the type of misspecification, and the underlying causal hypothesis (Baldwin, 1986).

Heise (1969) refers to phenomenon known as "theory-trimming." The deletion of causal linkages or factor loadings is often

dictated by the derived estimates of these parameters. Those estimates which are statistically negligible are simply assumed to be zero and are deleted from the model. Theory-trimming is completely assessed by the ability to reproduce the sample input matrix; theory-trimming becomes then a form of exploratory analysis in which previously hypothesized relationships are altered by statistical considerations of the data.

Conflicting viewpoints exist on the validity and permissible extent of theory-trimming. Cliff (1983), for example, refers to the "unreliability of hindsight." That is, there is a tendency to treat ex post facto analysis as confirmatory by establishing a theoretical model and testing its fit on a body of data. If this model proves a poor fit, it is modified and consequently "accepted." This modification procedure is an exploratory one however, not confirmatory, so that the researcher seemingly switches the purpose for research during the course of the analysis. Modification and acceptance of the model does not imply an improved model in the confirmatory sense. The new model fits better because the researcher "fixed" it to fit the data.

One must remember however that the original model describes reality according to available theory. If a model is rejected, it may be due to a lack of available theory or contradictions in theory. There is a subjective point beyond which a lack of knowledge cannot be tolerated, and exploratory analyses simply become number-crunching exercises. Saris, dePijper, & Zegwaart

(1979) suggest making small modifications but adhering to the original specifications as closely as possible. Even in these cases, modification of the model should always be judged in light of theoretical implications.

Cliff (1983) hints at cross-validation as a possible compromise to the exploratory/confirmatory dichotomy. Cudeck & Browne (1983) suggest splitting the sample data into two half samples-- using the first half to fit the model in an exploratory sense and then using the second half to confirm the model. Thus acceptance of the model is based not on one set of data but on two sets. In this sense, the goal of analysis is to find one or more possible models in the exploratory stage and then cross-validating to examine which model performs optimally for the second sample. A cross-validation index can be calculated which is a measure of the discrepancy between the reproduced covariance matrix of sample set 1 (the calibration sample) and the covariance matrix of sample set 2 (the validation sample). The choice of the model with the greatest predictive validity is based on the smallest value of the cross-validation index. Double cross-validation is also a possibility. The first wave of analysis proceeds as described above, the first sample set being used for calibration and the second for validation. The second wave of analysis then uses sample set 2 for calibration and sample set 1 for validation. After both stages of analysis are completed, the confirmed models can be compared.

Cudeck & Browne emphasize that the identification of one "best" model is not the objective of cross-validation. The goal of analysis should be to select one or more possible models and proceed to further study from that point.

Green (1977) expands the concept of interpretation beyond model-fitting by asking the following two questions: (1) how much statistical variation exists in the parameters; and (2) how sensitive is the model to changes in the parameters. If the intent of analysis is the understanding of cauasal dynamics, then the parameters must also be interpreted. Parameter values which change significantly from sample to sample may be too unstable to be interpreted realistically. The sensitivity of parameters relative to slight changes in model specification is also an indication of stability. There is a conceptual difference between parameter sensitivity and sampling variability which must be distinguished. Sampling variability is primarily dependent on the number of observations, and slight changes in parameters due to differences in sampling variability are usually not as serious as long as the sample is not exceedingly large or exceedingly small so as to make the parameters and goodness-of-fit statistics completely meaningless. Parameter sensitivity however is somewhat independent of N and reflects more on the nature of the model than on sample size.

LISREL-type structural equation modeling is a powerful statistical technique that seems especially appropriate for

social science variables which are inevitably complex and difficult to measure. The use of causal modeling methodologies implies attention to philosophical aspects of specification and interpretation as well as correct technical procedures. Granted, much is not known at the present time about the robustness of estimates with regard to violations of assumptions and modeling errors. Monte Carlo and analytical investigations are slowly making some of these issues more clear (see Boomsma, 1982; Baldwin, 1986; Ethington, 1985; and Muthen & Kaplan, 1985 for some current research on the robustness of LISREL maximum likelihood estimates). In the meantime,however, it is essential that the researcher proceed in LISREL analysis with as much technical and philosophical knowledge as possible.

Ethington, C. (1985). The robustness of LISREL estimates in structural equation models with categorical data. Unpublished dissertation, Virginia Polytechnical Institute.

Green, B. F. (1977). Parameter sensitivity in multivariate methods. Multivariate Behavioral Research, 12, 263-287.

Heise, D. R. (1969). Problems in path analysis and causal inference. In E. F. Borgatta (ed.), Sociological methodology, 1969. San Francisco: Jossey-Bass.

James, L. R. (1980). The unmeasured variables problem in path analysis. Journal of Applied Psychology, 65, 415-421.

James, L. R., Mulaik, S. A. & Brett, J. M. (1982). Causal analysis. Beverly Hills: Sage.

Joreskog, K. G. (1979). Structural equation models in the social sciences: specification, estimation, and testing. In K. G. Joreskog & D. Sorbom (Eds.), Advances in factor analysis and structural equation models. Cambridge: Abt.

Joreskog, K. G. & Sorbom, D. (1984). LISREL IV User's Guide. Mooresville, Indiana: Scientific Software, Inc.

Lomax, R. G. (1982). A guide to LISREL-type structural equation modeling. Behavior Research Methods & Instrumentation, 14, 1-8.

Lomax, R. G. (1983) A structural model of public and private schools. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal.

Lomax, R. G. (1986) The effects of measurement error in structural equation modeling. Ppaer presented at the Annual Meeting of the American Educational Research Association, San Francisco.

Miller, A. D. (1971). Logic of causal analysis. In H. M. Blalock (Ed.), Causal models in the social sciences. Chicago: Aldine.

Muthen, B. & Kaplan, D. (1986). A comparison of some methodologies for the factor analysis of non-normal Likert variables. British Journal of Mathematical and Statistical Psychology, 38, 171-189.

Namboordiri, N. K., Carter, L. R. & Blalock, H. M. (1975). Applied multivariate analysis and experimental design. New York: McGraw-Hill.

Saris, W. E., dePijper, W. M., & Zegwaart, P. (1979). Detection of specification errors in linear structural equation models. In K. F. Schuessler (Ed.), Sociological methodology, 1979. San Francisco: Jossey-Bass.

Strotz, R. H. & Wold, H. O. (1971). Recursive versus nonrecursive systems: an attempt at synthesis. In H. M. Blalock (Ed.), Causal models in the social sciences. Chicago: Aldine.

Sullivan, J. L. (1971). Multiple indicators and complex causal models. In H. M. Blalock (Ed.), Causal models in the social sciences. Chicago: Aldine.

Wilson, T. P. (1971). Critique of ordinal variables. In H. M. Blalock (Ed.), Causal models in the social sciences. Chicago: Aldine.

References

Baldwin, B. (1984). A causal model of the effects of maternal employment on adolescent achievement. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

Baldwin, B. (1986). The robustness of LISREL maximum likelihood estimates against model misspecification. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.

Bentler, P. M. (1980). Multivariate analysis with latent variables: causal modeling. Annual Review of Psychology, 31, 419-456.

Bentler. P. M. & Bonnett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin, 88, 588-606.

Blalock, H. M. (1982). Conceptualization amd measurement in the social sciences. Beverly Hills: Sage.

Borgatta, E. F. & Bohrnstedt, G. W. (1981). Level of measurement: once over again. In G. W. Bohrnstedt & E. F. Borgatta (Eds.), Social measurement: current issues. Beverly Hills: Sage.

Burt, R. S. (1981) A note on interpretational confounding of unobserved variables in structural equation models. In P. V. Marsden (Ed.), Linear models in social research. Beverly Hills: Sage.

Boomsma, A. (1983). On the robustness of LISREL (maximum likelihood estimation) against small sample size and non-normality. Amsterdan: Sociometric Research Foundation.

Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. Multivariate Behavorial Research, 18,115-126.

Costner , H. L. & Schoenberg, R. (1973) Diagnosing indicator ills in multiple indicator models. In A. S. Goldberger & O. D. Duncan (Eds.), Structural equation models in the social sciences. New York: Seminar Press.

Cudeck, R. & Browne, M. (1983). Cross-validation of covariance structures. Multivariate Behavorial Research, 18, 147-167.

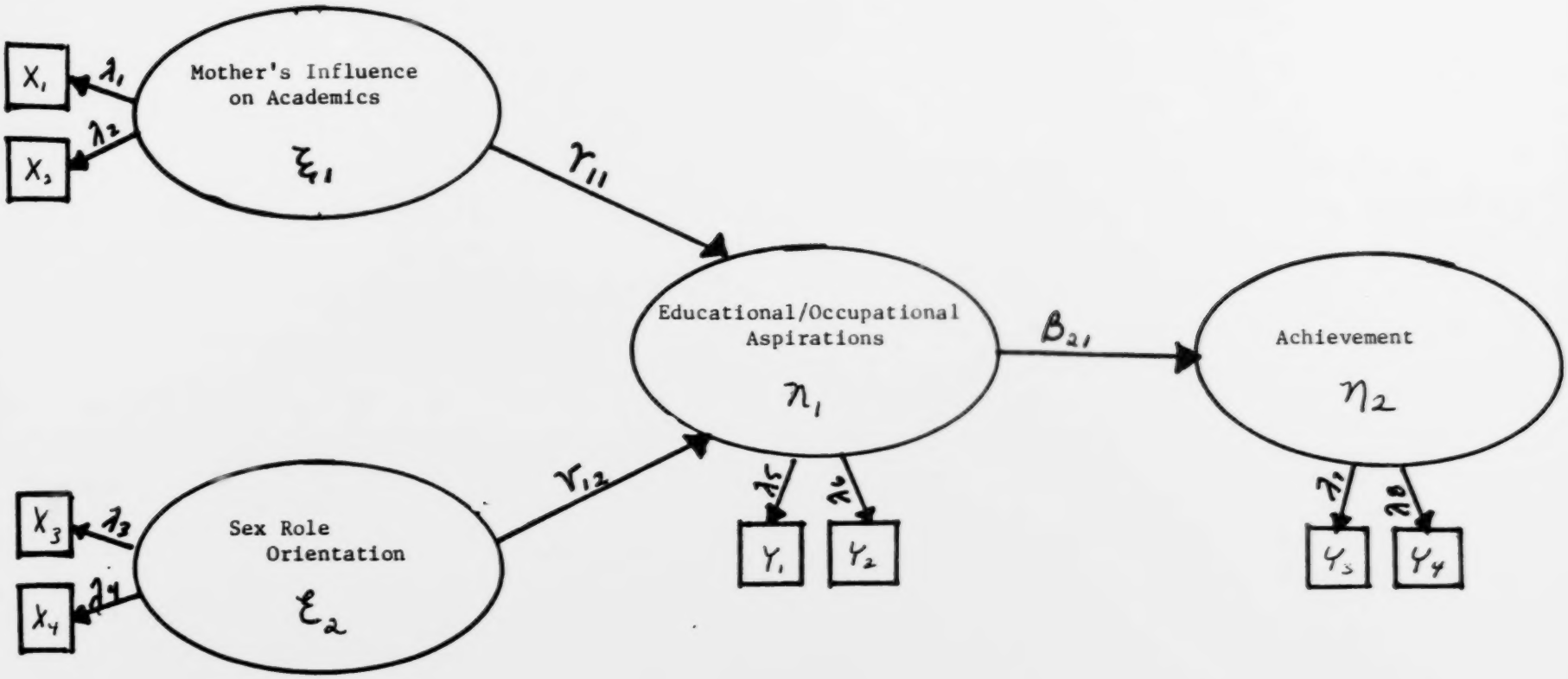


Figure 1. A Hypothetical Model of Mother-Daughter Role Modeling Relationships

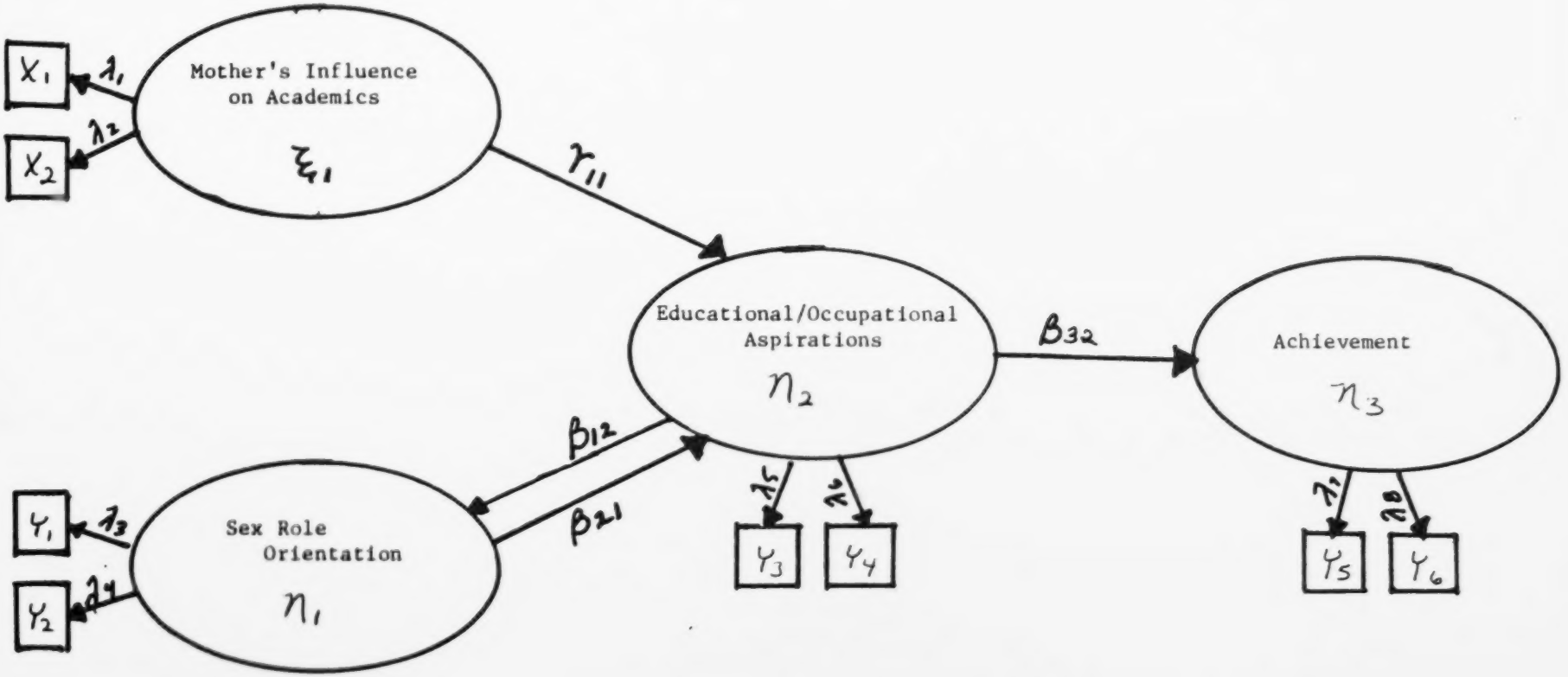


Figure 2. A Hypothetical Model of Mother-Daughter Role Modeling Relationships